Substrate recognition by the $tRNA^{His}$-guanylyltransferase: a kinetic investigation with model RNA substrates

A Senior Thesis

Presented in partial fulfillment of the requirements for graduation

with research distinction in Biochemistry in the undergraduate colleges of The Ohio State University

by

Lauren E. Duff

The Ohio State University

April 2015

Project Advisor: Dr. Jane E. Jackman, Department of Chemistry and Biochemistry

Abstract

All tRNAs undergo post-transcriptional modifications before being aminoacylated with a specific amino acid. In the case of $tRNA^{His}$, this includes the addition of a required guanosine at the -1 position (G_{-1}), required for histidyl-tRNA synthetase (HisRS) to recognize the tRNA and charge it with histidine. The $tRNA^{His}$ guanylyltransferase (Thg1) enzyme, with family members found in all domains of life, completes this reaction via its novel ability to add nucleotides in the 3'-5' direction in a chemically similar reaction to that of canonical 5'-3' polymerases. Eukaryotic Thg1 enzymes selectively recognize $tRNA^{His}$ for modification with G_{-1}, however the molecular basis for this observed substrate selectivity is not completely understood. Therefore, the long-term goal of this project is to elucidate the molecular features that participate in tRNA recognition by Thg1.

Insight into tRNA recognition by HisRS has been achieved by using small RNA stem-loop substrates and a similar approach is applied here to probe tRNA recognition by Thg1. The substrates to be tested are a substrate with a seven base pair stem that mimics the seven base pair acceptor stem of $tRNA^{His}$ and a 12 base pair substrate that mimics the acceptor stem and T-stem, which coaxially stack in full-length tRNA. The three Thg1 family enzymes tested for activity with these model RNAs are from *Methanobrevibacter smithii*, an archaeal species, *Bacillus thuringiensis*, a bacterial species, and *Saccharomyces cerevisiae*, a eukaryote. Here we report the results of kinetic assays with each of these purified enzymes and 5'-triphosphorylated substrates, all of which are analyzed using single-turnover assays that evaluate the rate of 3'-5' addition. Data from these assays will be used to evaluate the dependence of diverse Thg1 family enzymes on the length of the model RNA stem and its

potential for forming Watson-Crick vs. non-Watson-Crick base pairs during the 3’-5’ addition reaction.

Introduction

The genetic code contained within DNA is translated into functional proteins through two vitally important molecular biological processes: transcription and translation. During the course of transcription, DNA is copied into RNA, which is further processed by enzymes in cells. The mature RNA is utilized in translation to convert the genetic code into proteins. This latter process employs several different types of RNA including mRNA and tRNA. Messenger RNA (mRNA) is the direct copy of the DNA code that is read by ribosomes during translation. The adaptor molecule, transfer RNA (tRNA), is the molecule that allows ribosomes to read the mRNA. tRNA has an anticodon loop that contains a three base sequence which can pair with a complementary codon sequence in the mRNA and at the 3'-end it carries a specific amino acid, which is added to a growing polypeptide chain. These two processes need to be highly accurate and precise so that the genetic code can be correctly converted into functional proteins. Diseases and cancer can occur when transcription and translation fail to occur decorously, thus the enzymes involved in these pathways need to be exceedingly specific for their substrates.

To ensure that the DNA's message is appropriately translated, RNA molecules are highly processed. The main focus of this work is tRNA processing and modification. tRNA specifically undergoes numerous base modifications as well as aminoacylation, when an amino acid is added to the 3'-end by an aminoacyl-tRNA synthetase (aa-RS). In the case of $tRNA^{His}$ (as shown in Figure 1), HisRS charges it with the amino acid histidine because it recognizes the tRNA via its identity elements. One unique identity element contained in the structure of $tRNA^{His}$ is a guanosine at the -1 (G_{-1}) position. The $tRNA^{His}$ guanylyltransferase (Thg1) enzyme family has been found to complete this post-transcriptional modification in eukaryotes. This enzyme

family completes the transfer of the GTP nucleotide to the 5'-end of the tRNA via a novel 3'-5' three step reaction (Figure 2). This process occurs in a chemically similar mechanism to that of canonical, well-studied 5'-3' polymerases, but instead of the nucleophilic DNA strand attacking the incoming nucleotide, the incoming nucleotide acts as the nucleophile attacking the 5'-end of $tRNA^{His}$. The Thg1 mechanism first starts with an adenylylation of the 5'-monophosphate end $tRNA^{His}$, preparing it for the second step. The 3'-OH on the incoming GTP nucleotide attacks the activated 5'-end and releases AMP. The final step in this reaction catalyzed by Thg1 is pyrophosphate being cleaved from the triphosphorylated 5'-end. No viable Thg1 deletion strain can be produced which means that this enzyme is vital for growth, due to its role in $tRNA^{His}$ maturation and cell proliferation [1].

The Thg1 enzymes were first identified in *S. cerevisiae* and has since led to identification of many more Thg1 proteins in the three domains of life [3]. Thg1-like proteins (TLPs) were found using a BLAST search and are mainly found in Archaea and Bacteria, but also Eukarya [4]. Both Thg1 and TLPs use the same mechanism to add nucleotides, however they can have very different roles. In eukaryotes, Thg1 completes the addition of the single guanosine residue at the G_{-1} position across from an A_{73} residue in a non-Watson-Crick fashion. However in Bacteria, the G_{-1} is genomically encoded Watson-Crick across from C_{73}, which begs the question of why TLPs are present if they do not have to perform this essential reaction on $tRNA^{His}$. The genome of a eukaryotic slime mold, *D. discoideum* contains four TLPs (DdiTLPs); one of the four is a bona fide Thg1 ortholog, responsible for $tRNA^{His}$ maturation. Two out of the four, DdiTLP2 and DdiTLP3, play special roles in the mitochondria and the exact function for DdiTLP4 has yet to be found due to being unable to produce a genetically viable deletion strain. *In vitro* studies of

D. discoideum TLPs have shown that some TLPs have the capacity to catalyze 3'-5' nucleotide addition to 5'-truncated mitochondrial-tRNA substrates, activities consistent with 5' editing [3]. Understanding the specificity of the Thg1 enzyme family substrates is the goal of these studies in order to elucidate other possible substrates and functions that these enzymes can have in the cell.

Chapter 1: Make a conditional deletion strain for essential TLPs

The Thg1 function in eukaryotes has been extensively studied but the biological functions of TLPs found in the domains of Bacteria and Archaea are less understood [4]. In most archaeons and in nearly all bacterial species, the G_{-1} residue is genomically encoded and thus these organisms do not apparently require a Thg1 enzyme to add the G_{-1} residue. Due to the lack of necessity for Thg1 activity, Thg1/TLP family members may be required for other functions that utilize their 3'-5' addition reactions. Two functional characteristics have thus far been identified in TLPs; unlike yeast Thg1, bacterial/archaeal TLPs only add Watson-Crick base-paired nucleotides in the 3'-5' addition reaction and some can add missing 5'-nucleotides to tRNAs other than $tRNA^{His}$ [4]. To study these possible functions, a eukaryotic slime mold, *Dictyostelium discoideum,* was chosen because it encodes one Thg1 as well as three TLPs. DdiTLP3 and DdiTLP4 *in vitro* studies detected 5'-editing functions by adding missing 5'-nucleotides, however when attempting to create deletion strains of DdiThg1, DdiTLP3, and DdiTLP4 to observe the *in vivo* consequences, no viable colonies could be produced with traditional genetic methods. Thus, a different approach much be employed in order to

determine the *in vivo* functions of DdiTLP3 and DdiTLP4 and other potential functions for DdiThg1.

To solve this problem, a system needed to be fashioned so that a chromosomal deletion strain that could survive in the presence of a plasmid in which expression of the TLP can be controlled with an expression regulator. The tetracycline-Off (Tet-Off) expression system was chosen to control the amount of Thg1/TLP produced by the plasmid in the cells. In the Tet-Off system, the tetracycline-controlled trans-activator protein (tTA) is a fusion protein composed of the Tet-repressor DNA-binding protein (TetR) of the *E. coli* transposon Tn10 tetracycline resistance operon and domains that interact with the eukaryotic transcription machinery, which regulates expression of a target gene that is under transcriptional control of a tetracycline-responsive promoter element (TRE). The TRE is made up of Tet operator sequence fused to a promoter. In the absence of tetracycline, tTA binds to the TRE and activates transcription of the target gene. In the presence of tetracycline, tTA cannot bind to the TRE, and expression from the target gene remains inactive, as shown in Figure 3 [5].

The goal was to grow deletion strains of the three DdiTLPs that contain these plasmids with no tetracycline, allowing the cells to thrive due to the TLP being produced on the plasmid, and then increase the tetracycline concentration to visualize the effects of the loss of TLP function on the cell. This approach would be preferred because, in theory, the progression of loss of function on the cells could be seen over a period of time and the pressure on the cells could be reversed. To visualize the level of protein production, Western blots were performed on proteins recovered from cells that were grown with varying levels of tetracycline (0-30mM). On the Western blots, the recovered proteins should to migrate to the same size as their

respective pure protein to know that the correct protein was made by the plasmid. Also the expected levels of TLP production due to the Tet-OFF system should increase as the levels of Tet decrease as shown in Figure 4.

Results and Discussion

The results were varying for the three DdiTLPs. The Tet-controlled DdiThg1 and DdiTLP3 were apparently not present at high enough concentrations in the cell extract to be visualized. DdiTLP4, however, was readily detected by the TLP4-reactive antibody, as jusdged by the detection of a band that migrates in agreement with the expected molecular weight and that decreased in intensity when increasing concentration of tetracycline was included in the cultures. However, on all the blots tested, a large black smear or protein bands appeared between 45-66.7 kDa. This is assumed to be non-specific binding of the primary antibody to other proteins in the cell extracts or the antibodies binding to the membrane itself. A lager percent of milk helped decrease, but could not eliminate, this non-specific binding. To exclude the toxicity of tetracycline as a factor leading to the expected pattern, DdiTLP4 was exposed to αDdiThg1 and αDdiTLP3 primary antibodies. Importantly, these gels (in Figure 6) showed the relatively equal, barely detectable levels of DdiThg1 and DdiTLP3, respectively, present in the tet-DdiTLP4 and pure TLP4 protein samples. This means that the pattern related to tet-DdiTLP4 proteins was not due to the toxicity of tetracycline. From the blots, 3.0 µg/mL was the largest amount of tetracycline added that could produce still detectable DdiTLP4 and the next step would be to complete the *in vivo* studies.

Chapter 2: Testing small stem-loop RNAs for 3'-5' addition

To determine regions of $tRNA^{His}$ where other identity elements could be present and recognized by the Thg1 enzyme family, substrates needed to be made to target specific regions of the tRNA. Studies with aa-RS have shown that micro-RNA substrates can be aminoacylated when they mirror the acceptor stem of the tRNA [2]. A similar approach was used to create targeted substrates that mirror the 7 base pair acceptor stem (µHis7) and the 12 base pair acceptor stem plus the T-loop (µHis12) in order to discern if there is a minimal length of the coaxially stacked stem loops of $tRNA^{His}$ that Thg1/TLPs require for the addition reaction. The stem loop structures were designed from $tRNA^{His}$ and based on native folded tRNA structure (Figure 7). Along with different stem lengths, the nucleotide that is opposite of the added nucleotide was also considered, because in eukaryotes that is an A_{73} and in prokaryotes it is a C_{73}, which would predictably favor different base paired additions. Figure 8 shows all four of the substrates that were tested, including their specific sequences and the secondary structure of the micro-RNAs. The 73^{rd} nucleotide position on the full length $tRNA^{His}$ corresponds to the 32^{nd} nucleotide on the µHis12 substrates and the 22^{nd} position on the µHis7 substrates. To test the Watson-Crick preference of TLPs, GTP or UTP were used with the small stem-loop substrates to obtain the kinetic data of these two TLPs.

To isolate events at the active site of the enzyme without catalytic cycling, single-turnover conditions were utilized. In single turnover assays, the substrate was saturated with enzyme (E>>S) so that all of the substrate participated in the 'single turnover' and typically exhibited a single-exponential time course. Single turnover 3'-5' addition assays were performed to test for G_{-1} or U_{-1} addition to the micro-RNA substrates and thus characterize the

kinetic properties of TLP enzymes with either GTP or UTP. One TLP was chosen as a representative member from each domain of life: *M. smithii* TLP (MsTLP), *B. thuringiensis* TLP (BtTLP), and *S. cerevisiae* Thg1 (ScThg1). The kinetic analysis of ScThg1 was previously determined (Mess and Jackman, unpublished) and thus used for comparison to the TLPs.

Determining the kinetic features of these TLPs can help deduce what other substrates they can act upon as well as lead to finding other potential functions in the cell. For example, based on the substrates being utilized in this work, determining whether there is a size preference can lead to a conclusion that TLPs need to recognize an RNA substrate of a certain stem length in order to complete its reaction. Such a finding could lead to discovering another recognition sequence on $tRNA^{His}$ or on other RNA substrates. Another characteristic being examined is the preference for Watson-Crick (WC) versus non-Watson-Crick (non-WC) base pairing when catalyzing 3'-5' addition. It has been shown that TLPs generally only add NTPs in a WC fashion, however it is not known if under times of stress if these enzymes could add a non-WC base-paired nucleotide as Thg1 enzymes. Thus the goal of this project is to determine the kinetic preferences of size dependence and Watson-Crick base-pairing, which can have future implications of new substrates and functions for TLP enzymes.

Results

The activities of MsTLP and BtTLP were tested using the 5' γ-^{32}P-µHis substrates. To label the µHis RNAs, an in vitro transcription reaction was performed by T7 RNA polymerase in the presence of [γ-^{32}P] GTP which resulted in a single 5'γ-^{32}P triphosphate nucleotide that mimics an activated adenylylated $tRNA^{His}$ substrate from the first step of the Thg1 reaction. The single

turnover assays were accomplished by incubating each labeled substrate with 15 µM enzyme (MsTLP or BtTLP) in the presence of GTP or UTP. Reactions were then treated with RNase A to cleave the substrate RNAs after pyrimidines and EDTA to quench the reaction. When the incoming nucleotide attacks the labeled µHis RNA, the 5'-label is released as labeled pyrophosphate ($^{32}PP_i$), which can also decay over time to labeled inorganic phosphate ($^{32}P_i$) (see Figure 9).

The single turnover nucleotide addition assays were visualized (Figure 10) and percent product formation was measured by quantification of the released pyrophosphate (and inorganic phosphate) (P*Pi and Pi*). Figure 11 shows the general preference for TLPs to act upon the longer stem-loops as compared to the shorter ones. As summarized in Table 1, both TLPs could act upon both µHis12 substrates but neither exhibited any detectable activity with the µHis7 A_{22} substrates. Only BtTLP could act upon the µHis7 C_{22} substrate however it exhibited a slower rate than the µHis12 C_{32} substrate for both the G_{-1} and U_{-1} reactions. With the µHis12 C_{32} and µHis12 A_{32} substrates, BtTLP exhibited faster non-Watson-Crick base pairing reactions, opposite to what was expected for TLPs, since in the context of full length $tRNA^{His}$, BtTLP clearly prefers to catalyze a Watson-Crick base pairing. This effect was not evident in MsTLP which showed a more strict Watson-Crick preference.

Figure 12 demonstrates the kinetic preference for Watson-Crick base pairing by TLPs. Upon further testing, the K_D for each the nucleotides (GTP vs. UTP) add to the C_{32} containing substrate was assessed by performing single turnover assays with varied concentrations of GTP and UTP. For BtTLP, the K_D of the UTP reaction was 1.9 mM and the MsTLP K_D was larger than 1.0 as well. The catalytic efficiency of the reactions was found by dividing the averaged

observed rate of reaction by the K_D. For both TLPs, even though the U_{-1} reactions were faster, they were also less catalytically efficient than the G_{-1} reactions, by a factor of 11 fold for MsTLP and 4 fold for BtTLP, due to the significantly increased $K_{D,UTP}$ for the reaction catalyzed by each enzyme. Also Table 2 illustrates that BtTLP acted upon µHis12 C_{32} more efficiently than on the µHis7 C_{32} substrates. BtTLP was 22 fold more efficient than MsTLP at G_{-1} addition and 66 fold more than MsTLP at U_{-1} addition on µHis12 C_{32} substrates. As evident with both TLPs, the more efficient G_{-1} addition reaction also produced more product than the less efficient U_{-1} addition.

Discussion

Several conclusions can be made based on the data from the two different TLP experiments that were the focus of this thesis research. First, a tetracycline-controlled plasmid with the DdiTLP4 gene was the only plasmid that could produce DdiTLP protein in sufficient amounts to be detected by Western analysis. Second, based on the Western blot results, 3.0 mM tetracycline was the highest concentration that still produced detectable amounts of protein; however further testing would have to be done to determine if there is sufficient residual TLP4 expression under these conditions that would permit construction of a viable chromosomal ΔDdiTLP4 strain, in which case 0.3 mM tetracycline may have to be used as a maximum concentration. This approach would not be useful for DdiThg1 or DdiTLP3 because protein was not detected by the Western blot. The next steps to be taken with this project include creating a ΔDdiTLP4 strain, transform the tested plasmid into the cells, allow them to grow, then add tetracycline to the growth media, and determine the consequences of loss of TLP4 activity on the *D. discoideum* cells.

In the nucleotide addition assays with MsTLP and BtTLP, five general conclusions could be made. First, there was an apparent kinetic preference for the µHis12 substrates over the µHis7 ones, indicating that the TLPs react better when they recognize the longer stem-loop of the µHis12 substrate. This could be due to either an inherent length preference for the RNA substrate, or to the presences of a previous unknown identity element in the T-loop of the tRNA, would need to be tested by constructing new stem-loops that are longer than the µHis12 and intermediate of µHis7 and µHis12, and mutating the base-pairs of the µHis12 to determine if there is an unknown identity element because the mutation should render the substrate unrecognizable by the enzyme. Interestingly, this is the opposite of what was true for ScThg1, the eukaryotic Thg1 and reaction control. This is consistent with the previous observation that Thg1 enzymes perform different roles than TLPs, because TLPs act on different substrates, not just $tRNA^{His}$.

Second, the TLPs mostly exhibited faster rates for the Watson-Crick base pairing reactions, as was expected, however BtTLP exhibited faster absolute rates for the non-Watson-Crick U_{-1} across C_{32} base paring reactions than the Watson-Crick base pairing reactions, which was unexpected since in previous studies TLPs have shown a stronger preference for template-dependent 3'-5' nucleotide addition reactions [4]. This leads into the third conclusion that even though the absolute maximum rates may appear faster, the efficiency of the reaction played a larger role *in vitro* thus further demonstrating the kinetic preference of TLPs for Watson-Crick base pairing in 3'-5' nucleotide addition. Fourth, a greater efficiency rate also led to a greater production of products by the reaction. Lastly, BtTLP was generally more efficient at catalyzing template- dependent G_{-1} and U_{-1} addition to the 5'-ends of µHis RNAs than MsTLP. This set of

experiments with μHis substrates has shown that there is more to consider when searching for substrate specificity and identity elements in RNA for Thg1 enzymes.

Materials and Methods

***D. discoideum* cell lysis and protein retrieval.** The purification was performed at 4^{o}C, all buffers were prechilled and the Cell Cracker (by Balch and Rothman using 8.002mm ball) was cooled down before use. 5x10^{7} cells were collected and washed with ice cold Sorensen buffer (13mM KH_2PO_4, 4mM Na_2HPO_4, pH 6), spun at 800g for 3 min at 4^{o}C, and supernatant was removed. The pellet was then resuspended in 2 volumes of ice-cold mitochondrial lysis buffer (20mM Tris-HCl, pH 8, 0.2mM EDTA, 15% (V/V) glycerol, 1mM DTT, and protease inhibitor cocktail: 1mM PMSF, 2µg/mL of aprotinin and leupeptin, and 1 µg/mL pepstatin). The cell suspension was passed through the Cell Cracker 20 times. Then one at a time of the following: 414 µL of ddH_2O, 200 µL 20% (V/V) Triton X-100, and 1 µL of 4M KCl was added, followed by 20 strokes after each addition. The lysate was then vortexed for 10 sec every 5 min during a 15 min incubation on ice and centrifuged at 13,000g for 1 hr. Supernatant was transferred to a new tube, with an aliquot saved for a Bradford Assay, flash froze on dry ice for about a min, then put in the –80^{o}C freezer.

12% SDS-PAGE gel and sample preparation. The resolving gel consisted of 40% acrylamide (29:1), 1.5M Tris pH 8.8, 10% SDS, ddH_2O, 10% APS, and TEMED. The stacking gel consisted of 40% acrylamide (29:1), 1.0M Tris pH 6.8, 10% SDS, ddH_2O, 10% APS, and TEMED. Samples were diluted to desired concentration, 2 µL of dye (BME and sample buffer at a 1:10 ratio), boiled at 100^{o}C for 5-10 min. The gels were run at 180V for an hour, stained, shook for 20 min, and then destain solution was added and the gel was shook for an addition 2 hrs before being visualized.

Western blot. Samples (3 μg of purified protein, and 13 μg Thg1/TLP plus tetracycline proteins) were run on 12% SDS-PAGE gel then transferred to nitrocellulose (at 100V for 30 min). The membrane was stained with 0.1% (W/V) Ponceau S in 5% acetic acid, blocked for 1 hr with 5% milk in PBS at room temperature (RT), and washed with water. Concentrations of the primary antibodies had to be doubled in order to visualize the extracted TLPs versus the pure TLPs. Primary antibodies were used as follows: DdiThg1 1:100, DdiTLP3 1:250, and DdiTLP4 1:350 dilutions in 5% milk in PBS for titration proteins and 1:200, 1:500, and 1:700 dilutions in 5% milk in PBS, respectfully, for pure proteins. Incubate at RT for 4 hrs and then wash with water. Secondary antibody anti-rabbit IGg peroxidase (Sigma) was used at 1:13,000 dilution in 5% milk in PBS and incubated for 1 hr at RT. After the incubation, the membrane was washed 3 times for 5 min each with PBS + 0.005% Tween. Detection was done with ECL Plus Western Blotting Detection System (GE).

MsTLP and BtTLP Protein purification. Plasmids encoding MsTLP and BtTLP were transformed into *E. coli* strain BL21 DE3 pLysS, chosen for their capability of protein over-expression. Cultures were grown and proteins were purified by immobilized metal-ion affinity chromatography (IMAC) on TALON resin (Clontech) as described previously [3]. Resulting proteins were dialyzed in 50% glycerol for storage and assessed for purity (≥90%) by SDS-PAGE and stored at -20^{o}C. Protein concentrations were determined using the BioRad protein assay.

***In vitro* transcription of 5'-^{32}P-µHis substrates.** Each plasmid included a T7 promoter allowing for transcription as needed by the DNA-dependent RNA polymerase, T7 RNA polymerase in the presence of NTPs. DNA is incubated with T7 and done in the presences of [γ-^{32}P]GTP to generate 5'- γ ^{32}pppµHis substrates and for 1-2 hours at 37^{o}C, then treated with DNase I (RNase free) to break down the DNA into dNTPs, which can be separated from the intact RNA helices. After 30 minutes, substrates were purified by gel electrophoresis run on 10% polyacrylamide 4M urea gel. The bands were visualized by UV shadowing and cut out. The RNA was eluted from the gel with an RNA elution buffer. The substrates were PCA purified, ethanol precipitated, and resuspended in Tris/EDTA pH 7.5. Substrates were stored at -20^{o}C.

Single turnover 3'-5' addition assay. For these assays, BtTLP and MsTLP enzyme concentrations were held at 15 µM. Reactions were started by the addition of enzyme and carried out at room temperature in assay buffer containing 25 mM HEPES (pH 7.5), 10 mM $MgCl_2$, 3 mM DTT, 125 mM NaCl, 0.2 mg/mL BSA, GTP or UTP ranging from 0.1-10.0mM, and 5'-^{32}p*ppµHis substrate at 200cpm/µL (about 90nM). At indicated times, the reaction was stopped by removal of an aliquot (2 µL) to a new tube containing 0.5 µL of 0.5 M EDTA and 0.5 µL of 10 mg/mL RNaseA, and incubated to allow digestion for 10 min at 50°C. The reactions were treated with 1µL of 10% TCA, kept on ice for 5 min, and centrifuged for 5 min at 4°C to precipitate any unreacted RNA. 3 µL of each reaction were spotted on PEI cellulose-TLC plates and resolved in a 0.5M KPO_4:MeOH (80:20) solvent system. The plates were visualized and quantified using PhosphorImager and ImageQuant software.

Figures and Tables

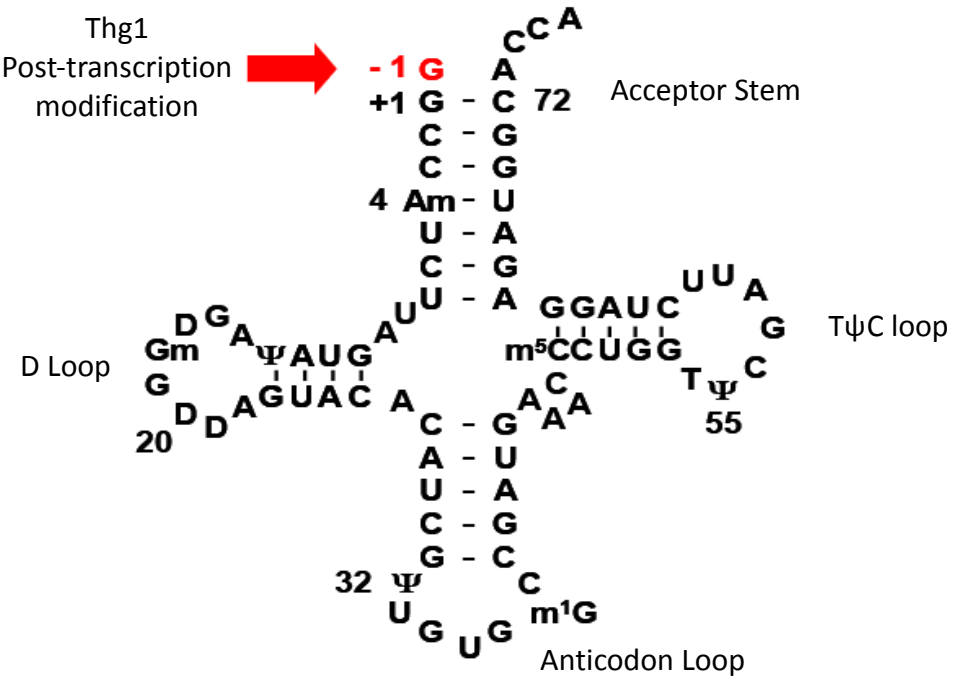

Figure 1. Mature tRNA^{His}. Conserved in all tRNA^{His} species G_{-1} is added post-transcriptionally in eukaryotes across A_{73} by Thg1 and genomically encoded in prokaryotes across C_{73}.

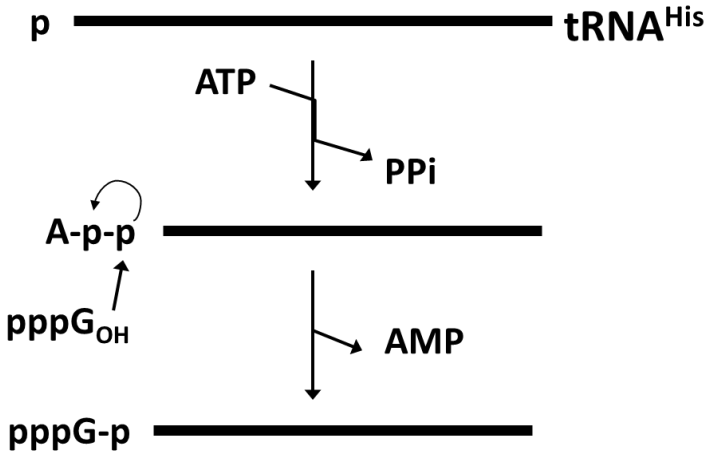

Figure 2. G_{-1} addition to tRNA^{His}. *In vivo*, this occurs via at least three chemical steps, requiring adenylylation, guanylyltransfer, and pyrophosphatase activities.

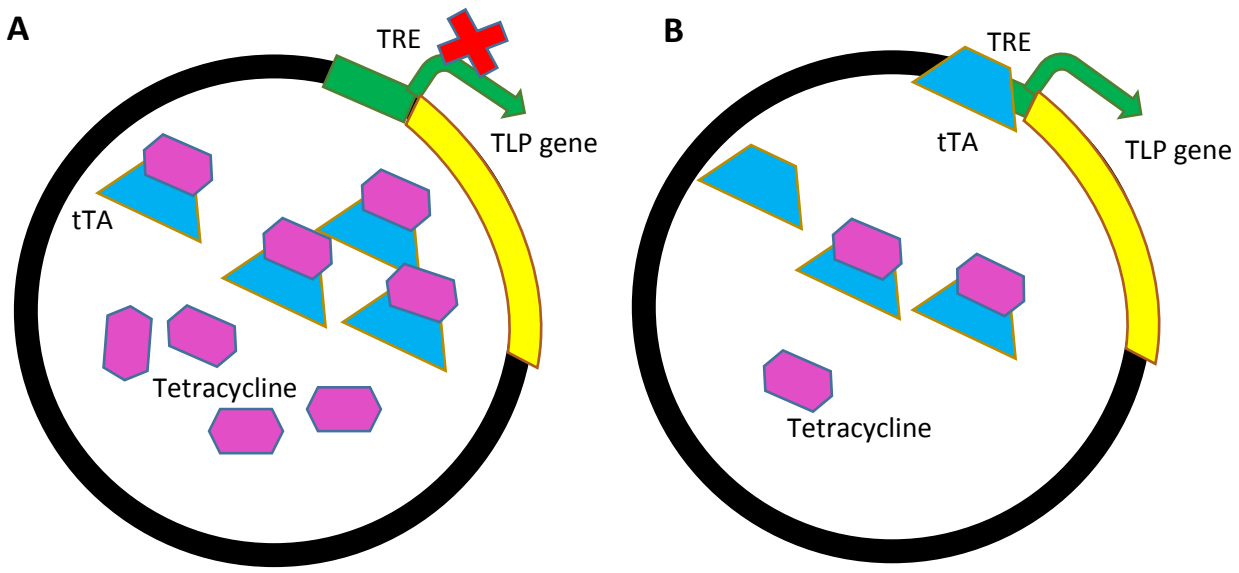

Figure 3. The Tet-OFF system controls the expression of the TLP gene. A) When tetracycline is present in high concentrations, there is no transcription due to binding of tetracycline to the tTA protein. B) When tetracycline is present in medium to low concentrations, there are varying degrees of expression because more tTA can access the TRE. tTA =Tet-controlled trans-activator protein, TRE= Tet-responsive promoter element.

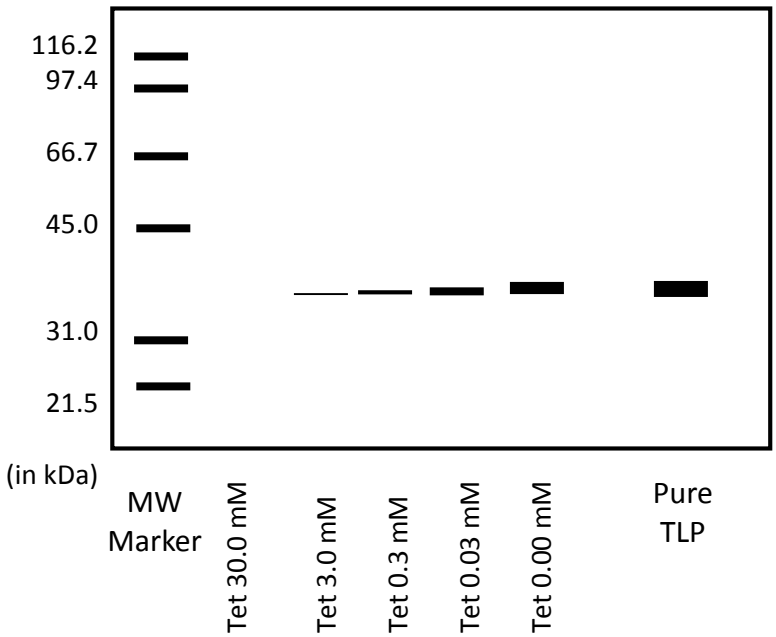

Figure 4. Experimental setup for Western Blots and expected outcome. As tetracycline is diluted, more protein is expected to have been produced by the plasmids. The pure TLP is used as a positive control and as indicator of where the Tet-controlled TLPs should migrate.

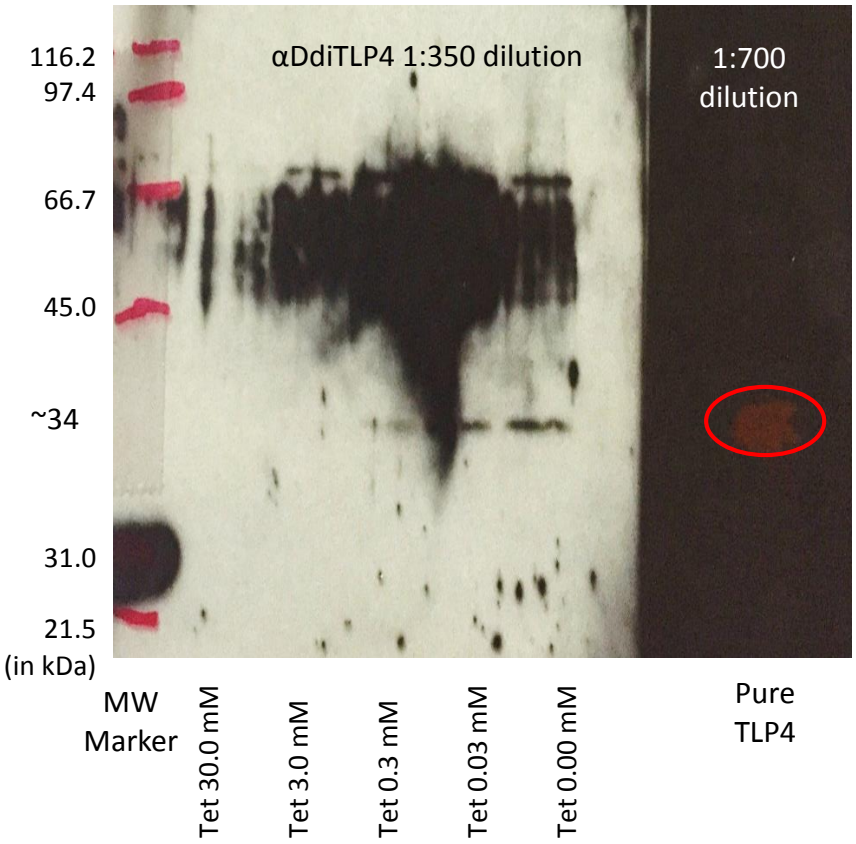

Figure 5. TLP4 set of proteins Western Blot. Samples (3 µg of pure TLP4, and 13 µg of TLP4 from cells grown with varying amounts of tetracycline) were run on 12% SDS-PAGE gel then transferred to nitrocellulose. The membrane was stained with 0.1% (W/V) Ponceau S in 5% acetic acid, blocked for 1 hr with 5% milk in PBS at room temperature (RT), and washed with water. Primary α-DdiTLP4 was a 1:350 dilution in 5% milk in PBS for titration proteins and 1:700 dilution in 5% milk in PBS for pure TLP4. Incubated at RT for 4 hrs and then wash with water. Secondary antibody α-rabbit IGg peroxidase (Sigma) was used at 1:13,000 dilution in 5% milk in PBS and incubated for 1 hr at RT. After the incubation, the membrane was washed 3 times for 5 min each with PBS + 0.005% Tween. Detection was performed with ECL Plus Western Blotting Detection System (GE).

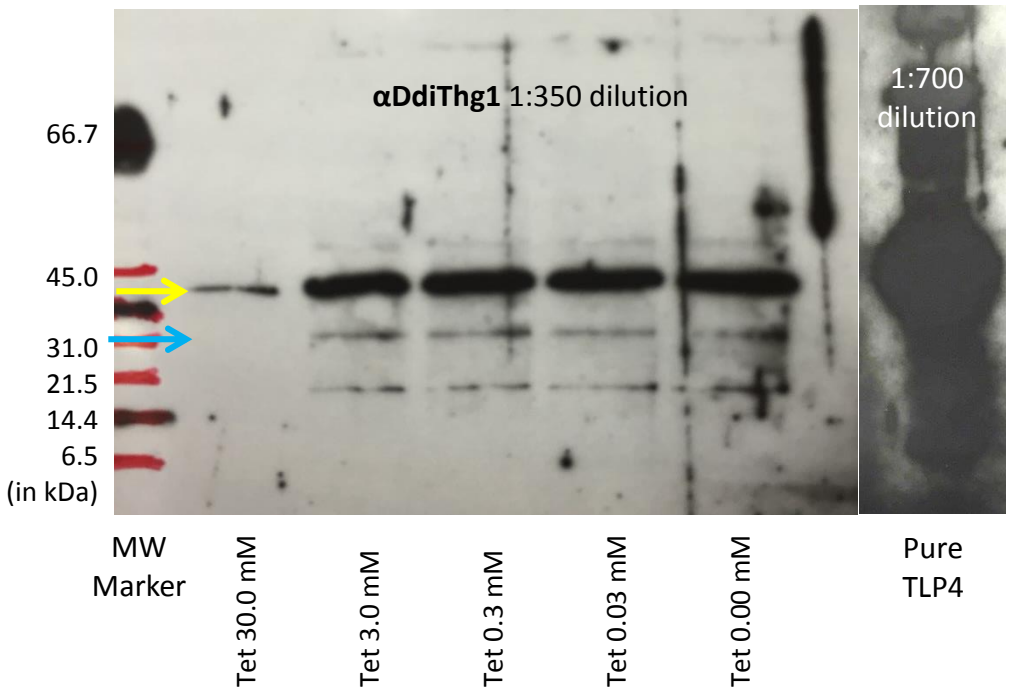

Figure 6. DdiTLP4 blot with αDdiThg1. On this blot of tet-DdiTLP4 proteins, αDdiThg1 primary antibody was used instead of αDdiTLP4 primary antibody to determine if Thg1 was present at low concentrations, which would prove that tetracycline is not toxic to the cells. DdiTLP4 has a molecular weight of 34 kDa (pointed out by the yellow arrow) and DdiThg1 has a molecular weight of 30 kDa (pointed out by the blue arrow). The smaller, equal sized Thg1 bands on the blot show that the large amounts of tetracycline in the cells are not toxic to them and thus do not influence the expected outcome.

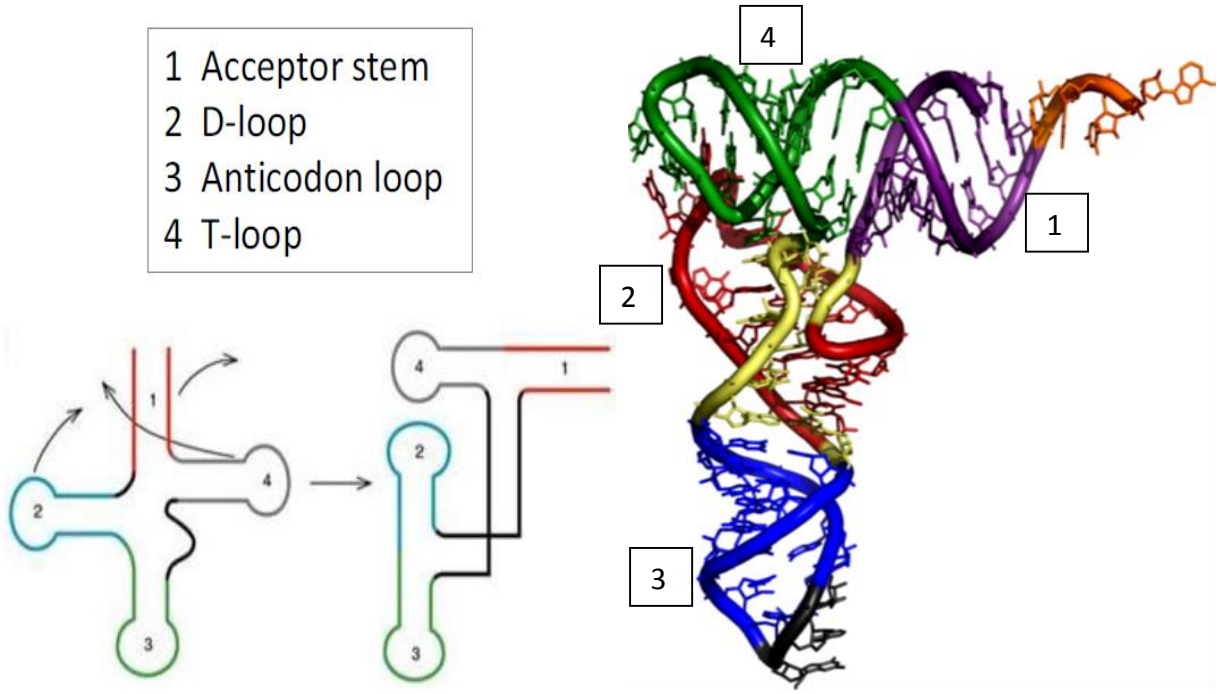

Figure 7. Showing how the different stems and loops of a tRNA are arranged in the folded tertiary structure. In the tertiary structure, the acceptor stem (1) and T-loop (4) are coaxially stacked.

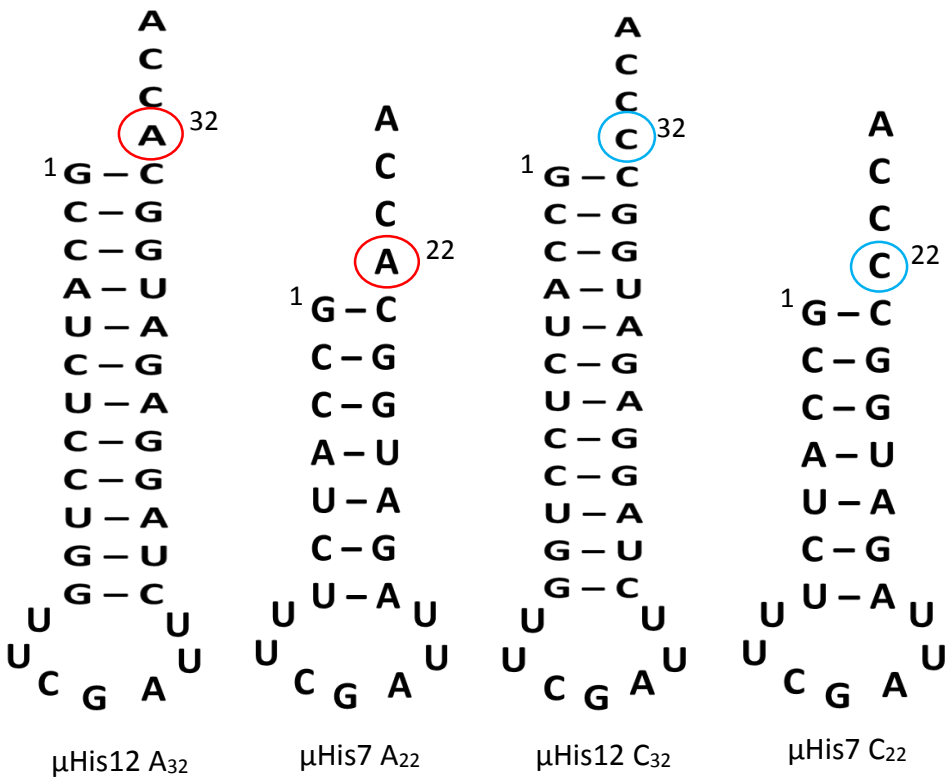

Figure 8. Single turnover assay substrates. Small RNA stem loops mimic the acceptor stem (µHis7) and the coaxially stacked acceptor stem and TψC loop (µHis12) of *S. cerevisiae* $tRNA^{His}$.

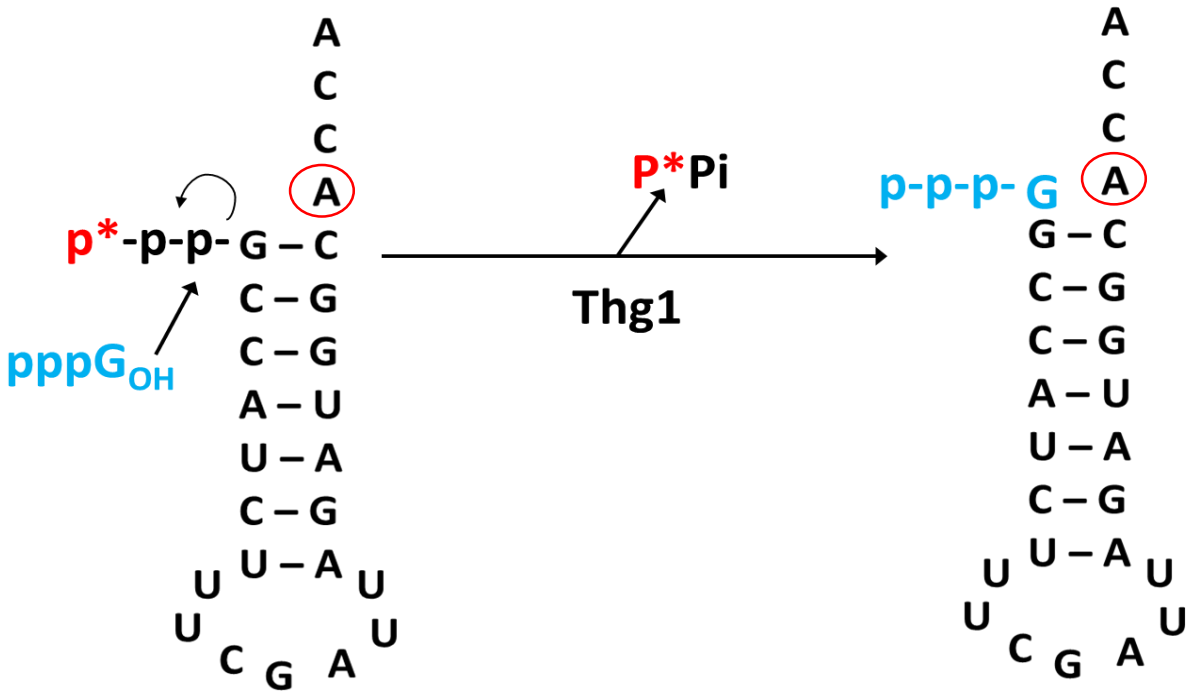

Figure 9. Thg1 Activity Assay. Labeled small RNA substrates were created by *in vitro* transcription in the presence of γ-^{32}P GTP for use in Thg1 activity assay; visualization of Thg1 catalytic activity is made possible by measuring release of PPi that occurs upon addition of the G_{-1} nucleotide, using TLC.

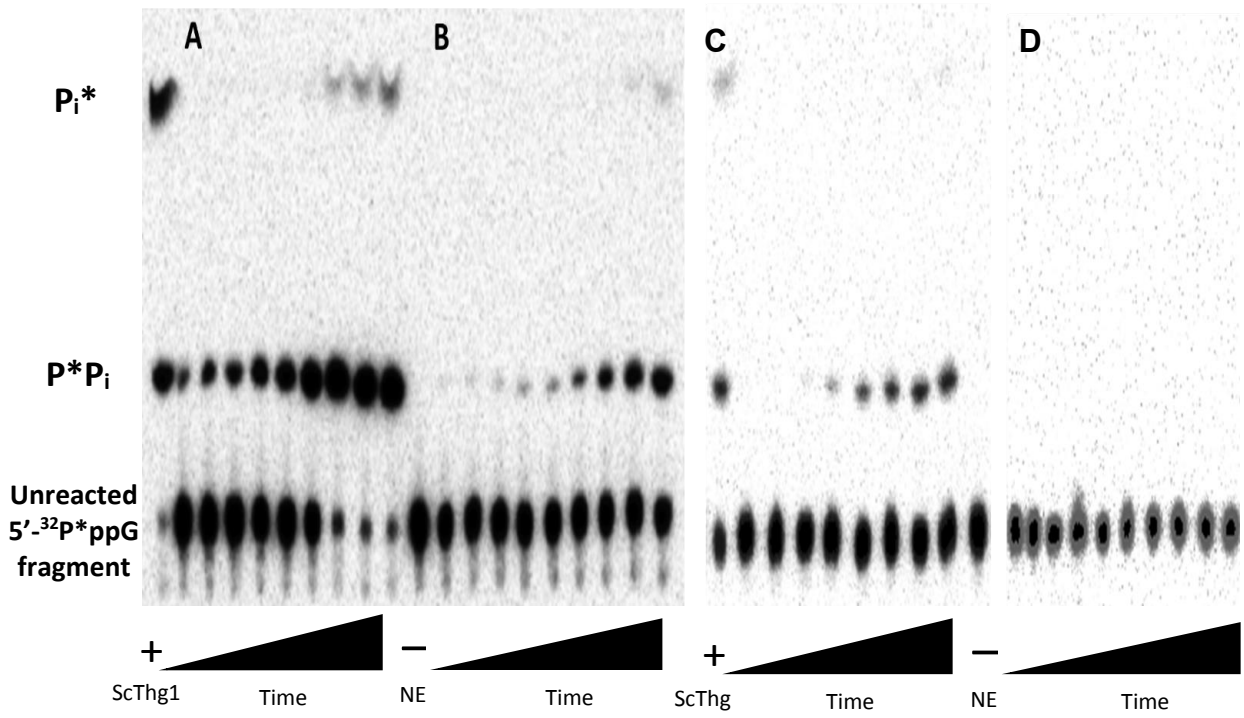

Figure 10. Representative single turnover Assays. Assay conditions: 25mM HEPES pH 7.5, 10mM $MgCl_2$, 3mM DTT, 125mM NaCl, 0.2mg/mL BSA, 1 mM NTP (GTP or UTP), ɣ-G tRNA substrate 200cpm/µL, and 15µM Thg1 or TLP enzyme. Reactions were initiated by addition of enzyme, quenched at the indicated time points by addition of 0.5µL EDTA and 0.5µL RNaseA, purified by trichloroacetic acid precipitation, and aliquots from each reaction were resolved by PEI-cellulose TLC in a 0.5 M potassium phosphate, pH 6.3:MeOH solvent system (80:20, v/v). NE = No Enzyme. A) µHis12 C_{32} + 1mM GTP, B) µHis7 C_{32} + 1mM GTP, C) µHis12 A_{32} + 1mM GTP, D) µHis7 A_{32} + 1mM GTP

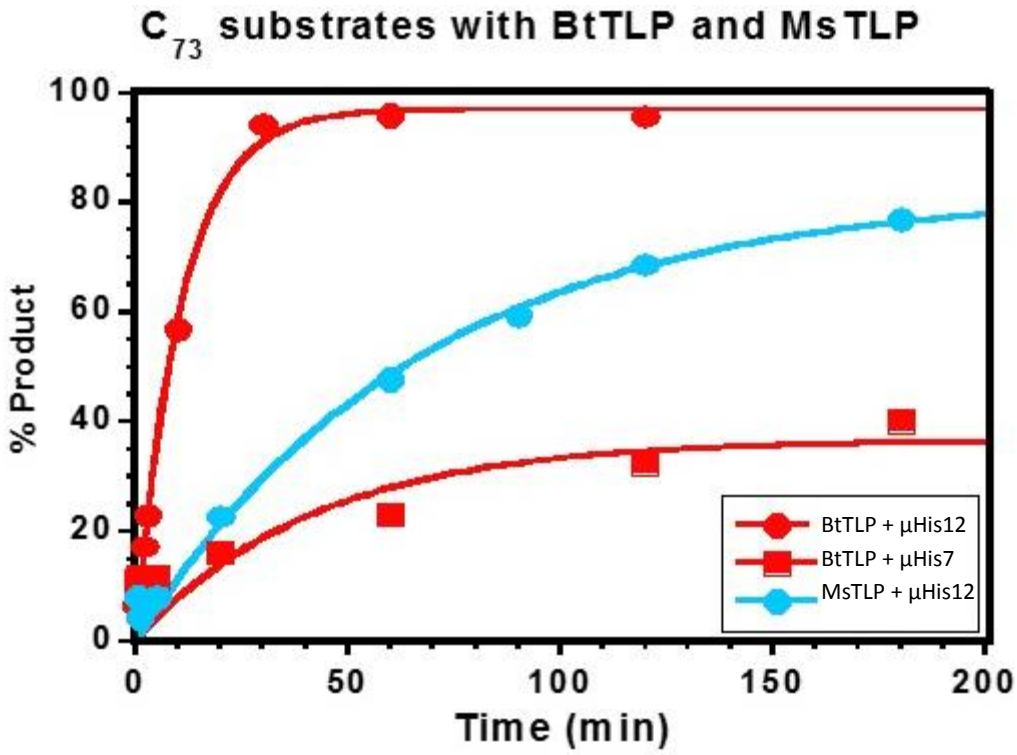

Figure 11. TLPs have a kinetic preference for longer stem-loops. MsTLP and BtTLP exhibited a kinetic preference for both the C_{32} and A_{32} substrates with faster rates and more product formation than the µHis7 C_{22} and A_{22} substrates.

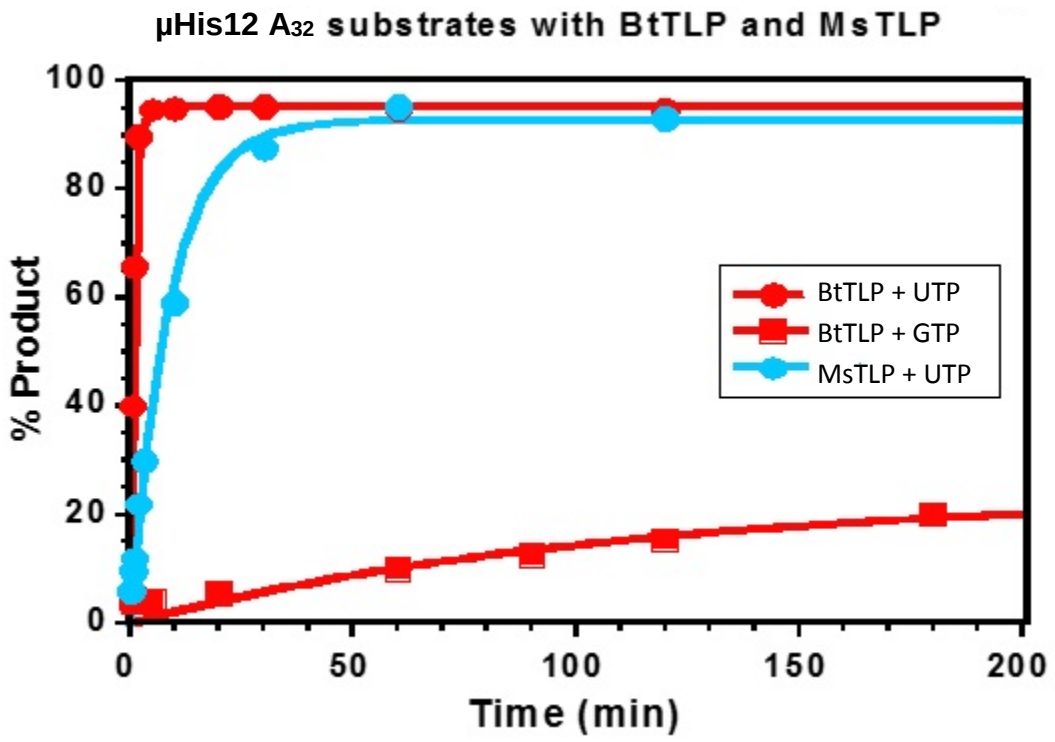

Figure 12. TLPs have a kinetic preference for Watson-Crick base pairing. MsTLP and BtTLP both show a kinetic preference for Watson-Crick base pairing with greater efficiency than non-Watson-Crick base pairing and generally more product formation, as evident in Table 3.

Enzyme	Substrate	1mM GTP $k_{Obs}(min^{-1})$	1mM UTP $k_{Obs}(min^{-1})$
S. cerevisiae* Thg1	µHis12 A_{22}	0.02	0.009
	µHis7 A_{22}	0.03	0.01
	µHis12 C_{32}	0.14	0.014
	µHis7 C_{22}	1.4	0.008
***M. smithii* TLP**	µHis12 A_{22}	**Not Detectable**	0.115 ± 0.008
	µHis7 A_{22}	**Not Detectable**	**Not Detectable**
	µHis12 C_{32}	0.022 ± 0.007	0.019 ± 0.008
	µHis7 C_{22}	**Not Detectable**	**Not Detectable**
***B. thuringiensis* TLP**	µHis12 A_{32}	0.005 ± 0.006	1.16 ± 0.04
	µHis7 A_{22}	**Not Detectable**	**Not Detectable**
	µHis12 C_{32}	0.5 ± 0.2	2.4 ± 1.6
	µHis7 C_{22}	0.006 ± 0.002	0.006 ± 0.002

Table 1. Observed rates of reaction for substrates tested. TLPs show different size dependence than positive control ScThg1, with TLPs preferring the longer substrates and ScThg1 the shorter. TLPs are able to complete non-Watson-Crick base pairing reactions and seem to have faster rates than the Watson-Crick reactions, however Table 2 illustrates that non-Watson-Crick reactions are less efficient. *Data obtained from Krishna Patel.

MsTLP

Substrate	µHis12, 1mM GTP	µHis12, 1mM UTP
k_{Obs} (min^{-1})	0.022 ± 0.007	0.019 ± 0.008
K_D (mM)	≤0.1	>1.0
k_{Obs}/K_D ($M^{-1}s^{-1}$)	≥3.67	<0.316

BtTLP

Substrate	µHis12, 1mM GTP	µHis12, 1mM UTP	µHis7, 1mM GTP/UTP
k_{Obs} (min^{-1})	0.5 ± 0.2	2.4 ± 1.6	0.006 ± 0.002
K_D (mM)	≤0.1	1.9	≤0.1
k_{Obs}/K_D ($M^{-1}s^{-1}$)	≥83.33	21.05	1.0

Table 2. Efficiencies of the TLP reactions with C32 substrates. Upon further testing, the faster rates for non-Watson-Crick reactions also had higher dissociation constants and thus lead to a lower efficiency value.

MsTLP

	µHis12 (C_{32}), 1mM GTP	µHis12 (C_{32}), 1mM UTP
Average Total % Product	82.29	36.17
k_{Obs}/K_D ($M^{-1}s^{-1}$)	≥3.67	<0.316

BtTLP

	µHis12 (C_{32}), 1mM GTP	µHis12 (C_{32}), 1mM UTP	µHis7 (C_{32}), 1mM GTP	µHis7 (C_{32}), 1mM UTP
Average Total % Product	93.84	92.29	42.21	17.49
k_{Obs}/K_D ($M^{-1}s^{-1}$)	≥83.33	21.05	1.0	1.0

Table 3. Percent product formation in relation to efficiency. The more efficient reactions lead to greater product formation.

References

1. Jackman J.E., Gott J.M., Gray M.W. *Doing it in reverse: 3'-to-5' polymerization by the Thg1 superfamily*. RNA, 2012. 18(5) p. :886-899.
2. Francklyn, C., K. Musier-Forsyth, and P. Schimmel, *Small RNA helices as substrates for aminoacylation and their relationship to charging of transfer RNAs.* European Journal of Biochemistry, 1992. 206(2): p. 315-21.
3. Abad, M.G., Rao, B.S., and Jackman, J.E. *Template-dependent 3'-5' nucleotide addition is a shared feature of tRNAHis guanylyltransferase enzymes from multiple domains of life.* Proceedings of the National Academy of Sciences, 2010. 107(2): p. 674-679.
4. Abad M.G., Long Y., Willcox A., Gott J.M., Gray M.W., Jackman J.E. *A role for $tRNA^{His}$ guanylyltransferase (Thg1)-like proteins from Dictyostelium discoideum in mitochondrial 5'-tRNA editing*. RNA, 2011. 17(4): p. 613-623.
5. Baron U, Bujard H. *Tet repressor-based system for regulated gene expression in eukaryotic cells: Principles and advances*. Methods Enzymol, 2000. 327: p. 401-21.